Operationalization of Differences among Protocols for Manual Hippocampal Segmentation: Quantitative Data for EADC-ADNI Consensual Decisions on a Harmonized Protocol.

Marina Boccardi, Martina Bocchetta, Rossana Ganzola, Nicolas Robitaille, Alberto Redolfi, George Bartzokis, Richard Camicioli, John G. Csernansky, Mony J. de Leon, Leyla deToledo-Morrell, Ronald J. Killiany, Stéphane Lehéricy, Johannes Pantel, Jens C. Pruessner, Hilkka Soininen, Craig Watson, Simon Duchesne, Clifford R. Jack Jr, Giovanni B Frisoni.

LENITEM (Laboratory of Epidemiology, Neuroimaging and Telemedicine) IRCCS – S. Giovanni di Dio – Fatebenefratelli Brescia, Italy (MB, MBocch, RG, AR, GBF); Department of Radiology, Université Laval and Centre de Recherche Université Laval – Robert Giffard, Quebec City, Canada (NR, SD); Department of Psychiatry, David Geffen School of Medicine at UCLA, Los Angeles, CA (GB); Department of Biomedical Engineering, Centre for Neuroscience, University of Alberta, Edmonton, Alberta, Canada. (RC); Department of Psychiatry and Behavioral Sciences, Northwestern University Feinberg School of Medicine, Chicago, IL, USA (JGC); New York University School of Medicine, Center for Brain Health, New York, NY (MJdL); Dept. of Neurological Sciences, Rush University, Chicago, Illinois (LdTM); Department of Anatomy and Neurobiology, Boston University School of Medicine (RJK); Center for NeuroImaging Research - CENIR and Dept of Neuroradiology, Université Pierre et Marie Curie-Paris 6, Groupe Hospitalier Pitié-Salpêtrière, Paris, France (SL); Department of psychiatry and Psychoterapy, University of Frankfurt/Main, Germany (JP); McGill Centre for Studies in Aging, Department of Psychiatry, McGill University, Montreal, Quebec, Canada (JCP); Dept of Neurology, University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland (HS). Wayne State University School of Medicine, 8D-University Health Center, St. Antoine, Detroit, MI (CW); Department of Diagnostic Radiology, Mayo Clinic and Foundation, Rochester, MN (CJ)

Introduction

A harmonized procedure is required, since quantitative MRI is to be used for diagnosis and tracking of Alzheimer's disease (AD). A survey of segmentation protocols allowed the identification of anatomical sources of heterogeneity in volume estimates. To identify differences among various MRI-based hippocampal segmentation protocols and provide quantitative information supporting decisions for an international harmonized protocol.

Methods

We operationalized landmark differences among protocols into segmentation units (SUs). We computed intra- and inter-rater reliability for each SU for both expert and naïve tracers on 20 ADNI scans (4 subjects per degree of visual medial temporal atrophy that we computed according to Scheltens *et al*. 1992), and estimated the contribution of SUs to the volume difference between AD, MCI and control subjects on a sample of 77 subjects, computed for having sufficient power to detect the informative value of each SU about AD-related atrophy. The 20 subjects for the computation of ICCs were 8 healthy controls, 3 AD, and 9 MCI patients. The 77 subjects were 31 controls with normal CSF Aβ levels, 23 (subsequently converted) MCI, and 23 AD patients. All MCI and AD had abnormal CSF Aβ levels.

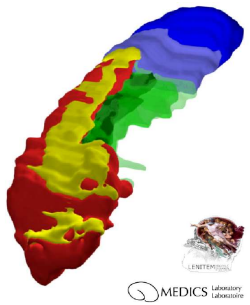
Results

We defined four SUs: Minimum Hippocampus (MinH), Alveus/Fimbria, Tail, and Subiculum. Reliability figures for SUs were >0.963 for intra-, and >0.905 for inter-rater, except Alveus/Fimbria (intra-rater: 0.863; inter-rater: 0.885). Noticeably, the intra-rater reliability of MinH and Alveus/Fimbria traced together (intra-rater: 0.993, 95% confidence interval: 0.983-0.997) was significantly higher than that of Alveus/Fimbria (95% confidence interval: 0.687-0.944). The intra-rater reliability re-computed by the naïve tracer provided analogous figures. The average volume difference between patients and controls was 538 mm^3, with MinH contributing to over 66% of this difference, Tail 27%, Alveus/Fimbria 12%, Subiculum over 5%. The SU volume differences between patients and controls were significant for all SUs except the Subiculum.

Intra- and inter-rater reliability of SUs computed on 20 ADNI subjects (4 by each degree of severity of hippocampal atrophy at the visual scale by Scheltens et al., 1992)

	Intra-rater	Inter-rater
MinH	0.992	0.974
Alveus/fimbria	0.863	0.885
MinH+Alveus/fimbria	0.993	0.973
Subiculum		
Oblique line	0.964	0.907
Morphology	0.981	0.937
Horizontal line	0.980	0.932
Tail		
Crura	0.998	0.937
Tail End	0.988	0.905

Rendering of the Maximum Hippocampus, composed by the sum of all SU. This is the largest possible hippocampus that can be segmented in theory considering the landmarks denoting the less restrictive definition for all sectors.



Segmentation Units volumes in controls, MCI and AD patients. Numbers denote mean volume (mm3) and standard deviation (in parentheses) of Segmentation Units. *p* denotes significance on t-test. Data are obtained on an ADNI sample of 77 subjects: 31 controls with normal CSF Aβ levels, 23 (subsequently converted) MCI, and 23 AD patients. All MCI and AD had abnormal CSF Aβ levels.

LEFT HIPPOCAMPUS

	Controls (n=31)	MCI (n=23)	AD (n=23)	*p* MCI vs CTR	*p* AD vs CTR	*p* MCI vs AD	*p* AD+MCI vs CTR
MinH	1467 (204)	1122 (263)	1023 (251)	<0,0005	<0,0005	0,199	<0,0005
Alveus/fimbria	248 (45)	232 (61)	200 (48)	0,269	<0,0005	0,055	0,01
Subiculum	243 (72)	220 (84)	213 (64)	0,279	0,118	0,754	0,121
Oblique line	196 (67)	178 (66)	176 (53)	0,338	0,262	0,936	0,207
Morphology	243 (72)	220 (84)	213 (64)	0,279	0,118	0,754	0,121
Horizontal line	234 (72)	210 (78)	211 (62)	0,243	0,221	0,957	0,153
Tail	485 (131)	383 (99)	353 (101)	0,003	<0,0005	0,307	<0,0005
Crura	190 (74)	177 (70)	146 (69)	0,538	0,034	0,14	0,101
End Tail	296 (120)	206 (76)	206 (86)	0,003	0,004	0,984	<0,0005
MaxHV	2443 (291)	1957 (348)	1788 (342)	<0,0005	<0,0005	0,105	<0,0005

RIGHT HIPPOCAMPUS

	Controls (n=31)	MCI (n=23)	AD (n=23)	*p* MCI vs CTR	*p* AD vs CTR	*p* MCI vs AD	*p* AD+MCI vs CTR
MinH	1462 (232)	1214 (247)	1061 (241)	<0,0005	<0,0005	0,039	<0,0005
Alveus/fimbria	255 (47)	258 (71)	225 (65)	0,84	0,05	0,103	0,35
Subiculum	225 (79)	208 (89)	184 (56)	0,459	0,042	0,294	0,109
Oblique line	181 (67)	167 (71)	150 (46)	0,455	0,059	0,334	0,122
Morphology	225 (79)	208 (89)	184 (56)	0,459	0,042	0,294	0,109
Horizontal line	220 (78)	203 (83)	182 (54)	0,459	0,053	0,309	0,117
Tail	487 (151)	349 (115)	349 (131)	0,001	0,001	0,999	<0,0005
Crura	187 (75)	169 (68)	140 (69)	0,37	0,025	0,17	0,058
End Tail	301 (120)	181 (113)	209 (110)	<0,0005	0,006	0,394	<0,0005
MaxHV	2429 (303)	2029 (372)	1820 (369)	<0,0005	<0,0005	0,062	<0,0005

Informative value of Segmentation Units for AD-related atrophy. Volumes (SD) are the same as in Table 2. Percent values denote the proportion of the SU compared to the total hippocampal volume (% of total hippo), the percent difference of the SU between groups (% diff), and the impact of the SU on the total volume difference between patients and controsl.

LEFT HIPPOCAMPUS

	Controls (n=31)	% of total hippo	MCI (n=23)	AD (n=23)	% diff MCI vs CTR	% MCI vs CTR (impact)	% diff AD vs CTR	% AD vs CTR (impact)
MinH	1467 (204)	60%	1122 (263)	1023 (251)	23,5%	71%	30%	68%
Alveus/fimbria	248 (45)	10%	232 (61)	200 (48)	6,5%	3%	19%	7%
Subiculum	243 (72)	10%	220 (84)	213 (64)	9,5%	5%	12%	4,5%
Oblique line	196 (67)	8%	178 (66)	176 (53)	9%	3,5%	10%	3%
Morphology	243 (72)	10%	220 (84)	213 (64)	9,5%	5%	12%	4,5%
Horizontal line	234 (72)	9%	210 (78)	211 (62)	10%	5%	10%	3,5%
Tail	485 (131)	20%	383 (99)	353 (101)	21%	21%	27%	20,5%
Crura	190 (74)	8%	177 (70)	146 (69)	6,5%	2,5%	23%	6,5%
End Tail	296 (120)	12%	206 (76)	206 (86)	30%	18,5%	30%	14%
MaxHV	2443 (291)	100%	1957 (348)	1788 (342)	20%	100%	27%	100%

RIGHT HIPPOCAMPUS

	Controls (n=31)	% of total hippo	MCI (n=23)	AD (n=23)	% diff MCI vs CTR	% MCI vs CTR (impact)	% diff AD vs CTR	% AD vs CTR (impact)
MinH	1462 (232)	60%	1214 (247)	1061 (241)	17%	62%	27%	66%
Alveus/fimbria	255 (47)	11%	258 (71)	225 (65)	-1%	-0,5%	12%	5%
Subiculum	225 (79)	9%	208 (89)	184 (56)	8%	4%	18%	6,5%
Oblique line	181 (67)	8%	167 (71)	150 (46)	8%	3,5%	17%	5%
Morphology	225 (79)	9%	208 (89)	184 (56)	8%	4%	18%	6,5%
Horizontal line	220 (78)	9%	203 (83)	182 (54)	7,5%	4%	17%	6%
Tail	487 (151)	20%	349 (115)	349 (131)	28,5%	34,5%	28,5%	22,5%
Crura	187 (75)	8%	169 (68)	140 (69)	10%	4,5%	25%	7,5%
End Tail	301 (120)	12%	181 (113)	209 (110)	40%	30%	31%	15%
MaxHV	2429 (303)	100%	2029 (372)	1820 (369)	16,5%	100%	25%	100%

Conclusions

Reliability of individual SUs and how informative they are in identifying AD-related atrophy will help define which SUs should be included in a harmonized protocol. Precise definition and dedicated attention to heterogeneity among anatomical landmarks should have an intrinsic harmonization value as illustrated by the high reliability demonstrated in the current study.